NASA TT F-9443

NASA TT F-9443

A RANDOM ITERATION PROCESS

Yu. V. Vorob'yev
(Leningrad)

Translation of "Sluchaynyy Iteratsionnyy Protsess".
Zhurnal Vychislitel'noy Matematiki i
Matematicheskoy Fiziki, Vol. 4, No. 6,
pp. 1088-1093, November-December, 1964.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
WASHINGTON D.C. JULY 1965

SCIENTIFIC REPORTS

UDC 518:5;7:949.12

A RANDOM ITERATION PROCESS

Yu. V. Vorob'yev
(Leningrad)

ABSTRACT

27722

A method of solving the problem Ax = f, where A is a linear operator, using random iteration processes is given.

The extensive use of program-controlled high-speed computers /1088* has lead to re-evaluations in the theory of approximation methods. The properties of these methods - such as the simplicity of their computational pattern, which reduces the programming volume, and their minimum load on the computer memory - have acquired great importance. The approach to an evaluation of the rate of convergence of the approximation process has also changed somewhat, and it is on this convergence rate that the number of computational operations directly depends.

As a rule, a high calculating speed makes a significant number of consecutive approximations possible, and asymptotic estimates play a decisive role in evaluating the convergence rate. For a savings in machine time, it is necessary that the error in the approximate solutions decrease sufficiently rapidly at the late stages of the

*Note: Numbers in the margin indicate pagination in the original foreign text.

approximation process. Significant deviations in the approximate solutions in the initial stage of the process, however, are of no essential importance, unless they affect the total extent of the computations.

Under these conditions, it seems to us, approximation methods which are based on random processes have a right to exist, i.e., not only the Monte-Carlo method in the strict sense - where the mathematical problem itself is replaced by an equivalent problem involving random quantities. This also includes the methods of solution, which do not change the formulation of the problem, in which the successive approximations represent the realization of some specially selected random process leading to a computation of the approximate solution with the required accuracy and with a considerable degree of probability.

One of these methods is proposed for the reader's consideration.

1. Iteration Processes

Let us examine the inhomogeneous linear equation

$$Ax = f,$$

where A is a linear, self-conjugate, positive definite, and restricted operator. Let us assume that we know the limits of its spectrum

$$m(x, x) \leqslant (Ax, x) \leqslant M(x, x).$$

The iteration processes are formed as follows. On the basis of an approximate solution of x_0, the successive approximations are

calculated from formula

$$x_{n+1} = x_n + \alpha_n r_n, \tag{1.1}$$

where $r_n = f - Ax_n$ is the residual of the n^{th} approximation.

If we designate the desired solution by x_*, the error of the n^{th} approximation $\eta_n = x_n - x_*$ is easily estimated. It follows that from Formula (1.1)

$$\eta_{n+1} = \eta_n - \alpha_n A\eta_n = (E - \alpha_n A)\,\eta_n$$

and

$$\eta_n = \prod_{k=0}^{n-1} (E - \alpha_k A)\,\eta_0 = \mathcal{P}_n(A)\,\eta_0,$$

after which the degree of error is readily estimated.

Using the spectral notation /1089

$$\eta_n = \int_m^M \mathcal{P}_n(\lambda)\, d\mathcal{E}_\lambda \eta_0,$$

we obtain

$$\|\eta_n\| \leqslant \max |\mathcal{P}_n(\lambda)|\, \|\eta_0\|, \tag{1.2}$$

where the greatest value of $|P_n(\lambda)|$ is taken on the segment [m, M] of the real axis

$$\mathcal{P}_n(\lambda) = (1 - \alpha_0\lambda)(1 - \alpha_1\lambda) \ldots (1 - \alpha_{n-1}\lambda).$$

How are the numbers α_k to be chosen? It is known from polynomial theory the estimate (1.2) is best if all the roots of the polynomial $P_n(\lambda)$ are located in the interval [m, M], so that

$$\frac{1}{M} \leqslant \alpha_k \leqslant \frac{1}{m}. \tag{1.3}$$

Various methods propose different ways of determining the numbers α_k and the polynomials $P_n(\lambda)$. An exposition of these iteration methods may be found in the monograph by D. K. Faddeyev and V. N. Faddeyeva (Ref. 1).

In the method of steepest descent, α_k is selected from the variational problem $\alpha_k = (r_k, r_k)/(Ar_k, r_k)$. Better convergence

is provided by processes in which sequences of orthogonal polynomials, with a certain distribution function which increases only in the segment [m, M], are utilized as the polynomials $P_n(\lambda)$. Two of these methods have been studied in detail — the method in which Chebyshev polynomials with the least deviation from zero are used as $P_n(\lambda)$ and the method of moments [see (Ref. 2), the minimum iteration method], in which the spectral function of operator A is the distribution function.

The latter group of methods no longer permits the use of the simplest algorithm described by formula (1.1).

We must resort to the construction of more complex algorithms by using recurrence formulas connecting the orthogonal polynomials.

We shall show that - by assuming α_k to equal a certain random variable α and by using the simplest algorithm (1.1) - we may produce an approximation sequence converging over the surface to the solution.

2. Random Polynomials

Let α be a random quantity and let $\alpha_0, \alpha_1, \ldots, \alpha_k \ldots$ be its values obtained as the result of independent trials. Let us consider the random polynomial

$$\mathcal{P}_n(\lambda) = (1 - \alpha_0\lambda)(1 - \alpha_1\lambda)\ldots(1 - \alpha_{n-1}\lambda).$$

Let $p_k = |1 - \alpha_k\lambda|$; then we have

$$\ln |\mathcal{P}_n| = \sum_{k=0}^{n-1} \ln p_k.$$

Since the random quantities $\ln p_k$ have one distribution function which is determined by that of the random quantity α, and since these quantities are obtained as the result of independent trials,

the conditions of the central limit theorem are fulfilled. Consequently, as n increases, the distribution law of ln $|\ P_n\ |$ approaches the normal one with a distribution density of

$$k(x) = \frac{1}{\sqrt{2\pi nD}} e^{-(x-na)^2/2nD},$$

where a and D are the mathematical expectation and the dispersion of the random quantities ln p_k.

By using the known formulas connecting the distribution functions, we find that the distribution function $|\ P_n(\lambda)\ |$ approaches the function

$$R(x) = \int_0^x k(\ln x)\,\frac{dx}{x} = \frac{1}{\sqrt{2\pi nD}} \int_0^x e^{-(\ln x - na)^2/2nD}\,\frac{dx}{x}, \quad x > 0,$$
$$R(x) \equiv 0, \quad x < 0.$$

We may now readily estimate the degree to which the random polynomial $P_n(\lambda)$ deviates from zero. Specifically, for any $\varepsilon > 0$ the probability that $|\ P_n(\lambda)\ | > \varepsilon$, when n is large, equals /1090

$$\mathrm{P}(|\mathscr{P}_n| > \varepsilon) \approx \frac{1}{\sqrt{2\pi nD}} \int_\varepsilon^\infty e^{-(\ln x - na)^2/2nD}\,\frac{dx}{x},$$

or, by using the Laplacian function

$$\Phi(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\,dt,$$

we finally obtain

$$\mathrm{P}(|\mathscr{P}_n| > \varepsilon) \approx 1 - \Phi\left(\frac{\ln \varepsilon - na}{\sqrt{2nD}}\right). \qquad (2.1)$$

In order that the probability [that the polynomial $P_n(\lambda)$ will deviate from zero] may approach zero as n increases, it is necessary and sufficient that the mathematical expectation

$$a = E\,[\ln|1 - \alpha\lambda|]$$

be less than zero.

3. Random Iteration Process

We must now choose f(x), the distribution density of the random

quantity α, so that the successive approximations x_n [determined in algorithm (1.1)] will converge in probability to the desired solution. First of all, by virtue of condition (1.3), f(x) differs from zero only in the segment [1/M, 1/m]. The mathematical expectation and dispersion of ln p are

$$a = \int_{1/M}^{1/m} \ln|\lambda x - 1| f(x)\, dx, \tag{3.1}$$

$$D = \int_{1/M}^{1/m} [\ln|\lambda x - 1| - a]^2 f(x)\, dx. \tag{3.2}$$

Let us discuss one of the possible methods for selecting the function f(x). Since we are forming a universal algorithm without assuming anything about the spectral distribution of operator A within [m, M], it is natural to require that the mathematical expectation a - which characterizes the convergence rate of the process - be independent of λ. This requirement leads us to the homogeneous, singular, integral equation

$$\frac{da}{d\lambda} = \int_{1/M}^{1/m} \frac{x f(x)}{\lambda x - 1}\, dx = 0,$$

whose solution, taking the normalization condition into account

$$\int_{1/M}^{1/m} f(x)\, dx = 1$$

gives the desired distribution density

$$f(x) = \frac{1}{\pi x \sqrt{(1 - mx)(Mx - 1)}}, \qquad m \leqslant x \leqslant M;$$
$$f(x) = 0 \quad \text{вне } [m, M].$$

In order not to interrupt the discussion, we shall postpone the not entirely elementary investigation of integrals (3.1) and (3.2) until the end of the paper. We shall merely present here the results of the computations. The mathematical expectation can be accurately computed:

$$a = -\ln \frac{\sqrt{M} + \sqrt{m}}{\sqrt{M} - \sqrt{m}}, \qquad (3.3)$$

while the dispersion is estimated /1091

$$D \leqslant \pi^2 + 8\sqrt{\frac{m}{M}} \ln 2 + O\left(\left(\frac{m}{M}\right)^{3/4}\right). \qquad (3.4)$$

Let us pass on to an evaluation of the convergence of approximate solutions. We shall use d to designate the greatest dispersion value D in segment [m, M]. It is also estimated by inequality (3.4). From formula (1.2) and (2.1), we can readily derive the following probability estimate of the error:

$$P\left(\frac{\|\eta_n\|}{\|\eta_0\|} > \varepsilon\right) \leqslant 1 - \Phi(X),$$

where

$$X = \frac{\sqrt{n}}{\sqrt{2d}} \ln\left\{\sqrt[n]{\varepsilon}\,\frac{\sqrt{M} + \sqrt{m}}{\sqrt{M} - \sqrt{m}}\right\}.$$

Defining the confidence coefficient as γ and the necessary accuracy as ε, we may find X_1 from the following equation by using probability integral tables:

$$1 - \Phi(X_1) = \gamma,$$

after which n, the number of necessary approximations, is found from the equality

$$X_1 = \frac{\sqrt{n}}{\sqrt{2d}} \ln\left\{\sqrt[n]{\varepsilon}\,\frac{\sqrt{M} + \sqrt{m}}{\sqrt{M} - \sqrt{m}}\right\}$$

Since $X_1 > 0$, approximate solutions - which are reliable to a certain extent - begin to be derived only beginning with values of n which satisfy the inequality

$$\left(\frac{\sqrt{M} - \sqrt{m}}{\sqrt{M} + \sqrt{m}}\right)^n \leqslant \varepsilon.$$

In practice, the solution is found as follows. Successive approximations are derived from algorithm (1.1):

$$x_{n+1} = x_n + \alpha_n (f - Ax_n).$$

The values of the random quantity $1/\alpha_n$, which has the distribution density

$$r(x) = \frac{1}{\pi\sqrt{(x-m)(M-x)}}, \qquad m \leqslant x \leqslant M,$$

may be computed from formula

$$\frac{1}{\alpha_n} = \frac{M-m}{2}\cos\pi\beta_n + \frac{M+m}{2},$$

where β_n are values of the random quantity β uniformly distributed in the interval [0, 1].

We would like to note a remarkable property of the function r(x) [see (Ref. 3)]. If in segment [m, M] the function p(x) is almost everywhere greater than zero and $P_n(x)$ is a sequence of orthogonal polynomials such that

$$\int_m^M P_k(x)\,P_j(x)\,p(x)\,dx = 0, \qquad k \neq j,$$

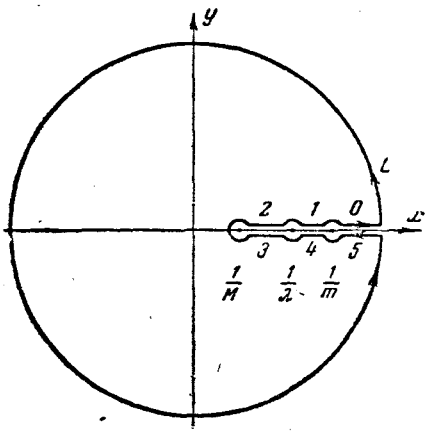
and if we introduce the distribution function of the zeros, $\psi_n(x)$ [which jumps by a positive value 1/n at every root of the polynomial $P_n(x)$], there is a single distribution limit function of the zeros $\psi(x)$, which is identical for all p(x) and towards which the sequence of $\psi_n(x)$ converges, and

$$\psi(x) = \int_m^\infty r(x)\,dx.$$

Thus, the limit function of the zero distribution of the orthogonal polynomials - including, of course, the Chebyshev polynomials - coincides with the zero distribution function of the random polynomials introduced by us.

4. Calculation of Integrals /1092

In order to compute the mathematical expectation (3.1)



$$a = \frac{1}{\pi} \int_{1/M}^{1/m} \frac{\ln|\lambda x - 1|}{x\sqrt{(1 - mx)(Mx - 1)}}\, dx$$

let us examine the function of the complex variable

$$F(z) = \frac{\ln(\lambda z - 1)}{z\sqrt{(1 - mz)(Mz - 1)}}$$

and contour L represented in the Figure. The branching points of F(z) are located at points 1/M, 1/λ, 1/m and the pole of the first order is located at zero. Fixing the branch of the many-valued function so that $F = \frac{\ln|\lambda x - 1|}{x\sqrt{(1 - mx)(Mx - 1)}}$ in region (1) of contour L, we shall obtain $F = \frac{\ln|\lambda x - 1| + i\pi}{x\sqrt{(1 - mx)(Mx - 1)}}$ in region (2); $F = -\frac{\ln|\lambda x - 1| + i\pi}{x\sqrt{(1 - mx)(Mx - 1)}}$ in region (3); $F = -\frac{\ln|\lambda x - 1| + 2i\pi}{x\sqrt{(1 - mx)(Mx - 1)}}$ in region (4); $F = +i\frac{\ln|\lambda x - 1|}{x\sqrt{(mx - 1)(Mx - 1)}}$ in region (0); and $F = +i\frac{\ln|\lambda x - 1| + 2i\pi}{x\sqrt{(mx - 1)(Mx - 1)}}$ in region (5). The remainder at zero equals π. Thus, we have

$$\int_L F(z)\, dz = 2\pi^2 i.$$

Letting the radius of the circle approach infinity and equating the real part of the integral to zero, we obtain

$$\int_{1/M}^{1/m} \frac{\ln|\lambda x - 1|}{x\sqrt{(1 - mx)(Mx - 1)}}\, dx = -\pi \int_{1/m}^{\infty} \frac{dx}{x\sqrt{(mx - 1)(Mx - 1)}}.$$

The last integral is tabulated, and a calculation of it results in formula (3.3).

A computation of the integral along the same contour from the function

$$G(z)=\frac{\ln^2\left[\frac{\sqrt{M}+\sqrt{m}}{\sqrt{M}-\sqrt{m}}(\lambda z-1)\right]}{z\sqrt{(1-mz)(Mz-1)}}$$

leads to the equality

$$D=\frac{1}{\pi}\int_{1/M}^{1/m}\frac{\ln^2\left[\frac{\sqrt{M}+\sqrt{m}}{\sqrt{M}-\sqrt{m}}|\lambda x-1|\right]}{x\sqrt{(1-mx)(Mx-1)}}dx=\pi\int_{1/\lambda}^{1/m}\frac{dx}{x\sqrt{(1-mx)(Mx-1)}}-$$

$$-2\int_{1/m}^{\infty}\frac{\ln\left[\frac{\sqrt{M}+\sqrt{m}}{\sqrt{M}-\sqrt{m}}|\lambda x-1|\right]}{x\sqrt{(mx-1)(Mx-1)}}dx.$$

The first integral on the right does not exceed π^2 for $m \leq \lambda \leq M$, /1093 while the second may be expanded in series in powers of $\sqrt{m/M}$. The estimate (3.4) is obtained from this.

Received December 23, 1963

REFERENCES

1. Faddeyev, D. K., and Faddeyeva, V. N. Linear Algebraic Computer Methods (Vychislitel'nyye metody lineynoy algebry). Fizmatgiz, Moscow, 1960.

2. Vorob'yev, Yu. V. The Moment Method in Applied Mathematics (Metod momentov v prikladnoy matematike). Fizmatgiz, Moscow, 1958.

3. Geronimus, Ya. L. The Theory of Orthogonal Polynomials (Teoriya ortogonal'nykh mnogochlenov). Gostekhizdat, Moscow-Leningrad, 1950.